METODY ILOŚCIOWE W BADANIACH EKONOMICZNYCH
X, 2009, str. 67 - 75

FORECASTING THE END-OF-THE-DAY REALIZED VARIANCE*

Małgorzata Doman
Katedra Matematyki Stosowanej, UE w Poznaniu
e-mail: malgorzata.doman@ue.poznan.pl
Ryszard Doman
Wydział Matematyki i Informatyki, AMU w Poznaniu
e-mail: ryd1oman@amu.edu.pl

Abstract. A large package of information is being reflected in stock prices during a short period after opening. Moreover, the start-of-the-day (morning) volatility has a strong impact on the price variability during all the day. In this connection, the question is whether the morning realized variance calculated as the sum of morning squared intraday returns can be useful in forecasting the daily realized variance (end-of-the-day volatility). In the paper, we apply three different methods of forecasting the daily realized variance for stocks quoted on the Warsaw Stock Exchange Our findings show that the morning realized variance provides valuable information that can be used in forecasting the daily realized variance.

Keywords: realized volatility, forecasts, ARFIMA, unobserved component model

INTRODUCTION

The accurate measurement, estimation and forecasting of stock market volatility is a crucial task in portfolio management, asset pricing (especially, derivatives pricing) and risk management. The most popular definition used in quantitative finance is that the price volatility is the variance of return conditional on the information available a period earlier. Thus, volatility is an unobservable variable and its measurement depends on a model. The most common approach to model the stock market volatility is to use a model from the GARCH family [Tsay 2002]. Another possible method is to calculate the so-called implied volatility [Poon

* This work was financed from the Polish science budget resources in the years 2007-2010 as the research project NN 111 1256 33.

2005]. The daily realized volatility or the realized variance, proposed by Andersen and Bollerslev [1998], is calculated as a sum of squared intraday returns from the day.

The intraday returns coming from different part of the day give different input into the daily realized variance. A very large package of information is being reflected in prices during a short period after opening the stock. Therefore, the start-of-the-day (morning) volatility has a strong impact on the return dynamics during all the day. In this connection, the question arises of whether the morning realized variance can be useful in forecasting the daily realized variance (end-of-the-day volatility).

The aim of the presented research is to assess to what extent the intraday data from the beginning of the day can explain and predict the end-of-the-day volatility. We use the realized variance as a volatility measure and show that the realized variance calculated for first few hours of the trading day (MORNING realized variance) is a useful predictor of the daily volatility. The idea of our approach is slightly similar to the one presented in [Frijns and Margaritis 2008] though the applied way of describing the realized variance dynamics is different. We apply 3 methods of forecasting the daily realized variance of the stock returns. They include a direct modeling as an ARFIMA process, as an ARFIMA process with MORNING realized variance as an additional explanatory variable, and by using an unobserved component model. The analysis is performed for four chosen stocks from the Warsaw Stock Exchange (AGORA, TVN, BZWBK and PEKAO). Our results show that the daily volatility forecasts based on the morning realized variance can be a useful tool in short term investments.

REALIZED VARIANCE AND MARKET MICROSTRUCTURE EFFECTS

In this paper we consider logarithmic percentage returns

$$R_t = 100(\ln P_t - \ln P_{t-1}), \tag{1}$$

where P_t is the closing price on trading day *t*.

The daily volatility σ_t^2 of a financial instrument is defined as the conditional variance of its daily return given the set of information Ω_{t-1} available on day $t-1$, i.e.

$$\sigma_t^2 = E((R_t - E(R_t \mid \Omega_{t-1}))^2 \mid \Omega_{t-1}) \tag{2}$$

Thus the volatility is an unobservable variable and (assuming that the conditional mean of a daily return equals zero) the square of a daily return is an unbiased estimator of it. However, there exist many empirical findings showing that this estimator is very noisy [Andersen and Bollerslev 1998].

In their paper devoted to forecasting abilities of GARCH models, Andersen and Bollerslev [1998] proposed another estimator of daily volatility: the sum of squared intraday returns, and called it the realized volatility (variance). We denote the daily realized variance by $\tilde{\sigma}^2_{[DAY]t}$. Thus when we consider D+1 intraday quotations, $p_{t,d}$, $d = 0,1,\ldots,D$, we have

$$\tilde{\sigma}^2_{[DAY]t} = \sum_{d=0}^{D} r^2_{t,d} , \qquad (3)$$

where $r_{t,d} = 100(\ln p_{t,d} - \ln p_{t,d-1})$, $p_{t,0}$ is the opening price, and $p_{t,-1} = p_{t-1,D}$ is the closing price on the day before.

THE MODELS

The realized volatility exhibits long memory and this phenomenon suggests a possibility of describing the realized volatility dynamics by using ARFIMA models [Granger and Joyeux 1980, Hosking 1981].

An ARFIMA process x_t may be defined by

$$\alpha(L)(1-L)^d(x_t - \mu) = \beta(L)\varepsilon_t , \qquad (4)$$

where ε_t is a white noise process, $\alpha(L)$ and $\beta(L)$ are the lag polynomials of order p (autoregressive) and q (moving average), respectively, and $(1-L)^d$ is a fractional differencing operator defined by the binomial expansion. If $|d| < 0.5$ and the roots of $\alpha(L)$ and $\beta(L)$ lie outside the unit circle, the process x_t is stationary and invertible. For $0 < d < 0.5$, x_t displays long memory property. ARFIMA models are the most common way of modeling and forecasting the daily realized volatility [Koopman et al. 2005, Doman 2006]. It is possible to extend the ARFIMA model by introducing additional explanatory variables into the model equation. In this case, the constant μ in (4) is replaced by

$$\mu_t = \mu + \sum_{i=1}^{n} c_i x_{i,t} . \qquad (5)$$

The idea of modeling the realized volatility by unobserved component (UC) models comes from the results of Barndorff-Nielsen and Shephard [2002]. They studied the statistical properties of the volatility estimate error, $\sigma_t^2 - \tilde{\sigma}_t^2$, where

$\tilde{\sigma}_t^2$ is the estimator of the realized volatility based on intraday returns, and suggested modeling the spot volatility as some continuous-time stochastic process, more precisely, as a sum of independent Ornstein-Uhlenbeck processes. According to their results, in such a situation the actual volatility corresponding to day intervals can be modeled as the sum of ARMA(1,1) processes.

Following Koopman et al. [2005], we use only one ARMA(1,1) process and the specification of a UC model given by the equations

$$\tilde{\sigma}_t^2 = \sigma_t^2 + u_t \quad u_t \sim N\left(0, \sigma_u^2\right), \tag{6}$$

$$\sigma_t^2 = \mu + \phi\left(\sigma_{t-1}^2 - \mu\right) + \theta\eta_{t-1} + \eta_t \quad \eta_t \sim N\left(0, \sigma_\eta^2\right) \tag{7}$$

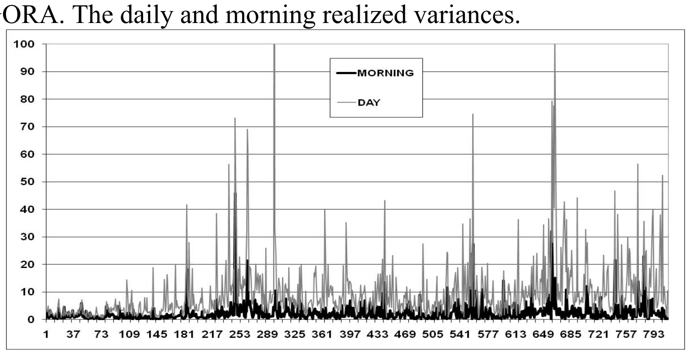
The estimated parameters are μ, ϕ, θ and σ_u^2. The variation of the error η_t is fixed. The model equations (6) and (7) can be represented in the state-space form, and parameter estimates as well as one-day- ahead volatility forecasts can be obtained by application of the Kalman filter [Durbin and Koopman 2002].

THE DATA

Our data set includes the intraday quotations of four companies listed on the Warsaw Stock Exchange. These are two media companies: AGORA and TVN, and two banks: BZWBK and PEKAO. The period under scrutiny is from June 1, 2005 to August 22, 2008. It is divided into the in-sample period (761 days) and the out-of sample period (50 days) for which the realized variance forecasts are calculated.

The realized variance is based on 5-minute returns. We consider the daily realized variance (DAY) and the MORNING realized variance corresponding to the period from the opening to 11:00. The MORNING realized variance is in average about 0.3 of the daily realized variance in each case. The plots in Figure 1 show the general impression of the dynamics of the considered realized variances.

Figure 1. AGORA. The daily and morning realized variances.



Source: Own calculation.

EMPIRICAL RESULTS

To answer the question about the usefulness of the morning volatility in prediction of all the day volatility, we applied 3 methods of modeling and forecasting the daily realized variance for the investigated stock returns. Besides the most common approach of modeling the daily realized variance as an ARFIMA process, we propose also to consider the ARFIMA process with the MORNING realized variance as an additional explanatory variable. The third of the considered possibilities is to use unobserved component models described in section 3.

It is a known phenomenon that the realized variance exhibits long memory [Andersen at al. 2001]. This effect is confirmed by our results. When considering ARFIMA models, the estimates of parameter d have values about 0.2 and all are significant (Table 2). The interesting observation is that, except the case of BZWBK, there is no ARMA part in the fitted models.

Table 1. Parameters estimates of fitted ARFIMA model

Parameters	AGORA	BZWBK	PEKAO	TVN
a_0	1.8420 (0.0795)	7.9670 (1.1218)	4.1132 (0.3405)	1.9007 (0.4540)
d	0.1673 (0.01485)	0.1346 (0.0574)	0.1939 (0.0295)	0.2009 (0.0162)
a_1		0.6099 (0.1812)		
b_1		0.4875 (0.1937)		

Source: Own calculation.

Table 2. Parameters estimates of fitted ARFIMA model with additional explanatory variable MORNING

Parameters	AGORA	BZWBK	PEKAO	TVN
a_0	0.6533 (0.1299)	3.3245 (0.8809)	2.2905 (0.2455)	1.8763 (0.3270)
d	0.1461 (0.0133)	0.2091 (0.0536)	0.1719 (0.0226)	0.1204 (0.0189)
a_1		0.8820 (0.1069)	-0.0051 (0.0424)	0.0797 (0.0469)
b_1		0.9081 (0.0789)		
c_1	1.1522 (0.1012)	1.2967 (0.0788)	1.1839 (0.0623)	1.3159 (0.0554)

Source: Own calculation.

Introducing into the ARFIMA model equation the additional explanatory variable MORNING, results in lowering the estimate of d (except BZWBK). The estimates of parameter c_1 are significant implying the importance of information contained in MORNING for the explanation of the daily realized variance dynamics (Table 3).

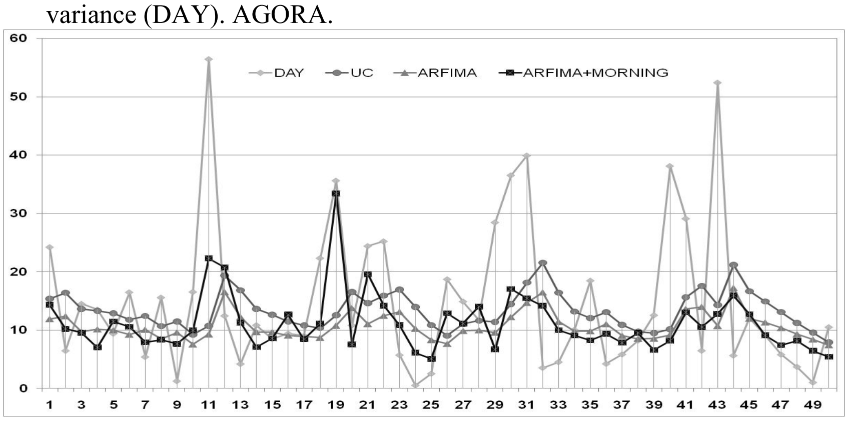
Although our earlier results [Doman and Doman 2004] show that UC models can be very useful in predicting the daily realized variance, the estimation results presented in Table 4 are not very impressive. It seems from the estimates of the variance parameters that this time the UC model turned out to not be so successful in capturing the realized variance dynamics.

Table 3. Parameters estimates of fitted UC models.

Parameters	AGORA	BZWBK	PEKAO	TVN
μ	9.4329 (1.1134)	13.0234 (0.7403)	7.9647 (0.5654)	10.7288 (1.2803)
ϕ	0.9912 (0.0523)	0.7703 (0.0582)	0.8392 (0.0487)	0.9000 (0.0305)
θ	-0.6685 (0.0903)	-0.4057 (0.0933)	-0.5013 (0.0885)	-0.6466 (0.0576)
σ_u^2	15.834 7.0362 35.634	18.676 12.676 27.517	5.6360 3.4562 9.1908	22.126 13.2819 36.8590
σ_η	11.0	8.0	5.0	10.0

Source: Own elaboration.

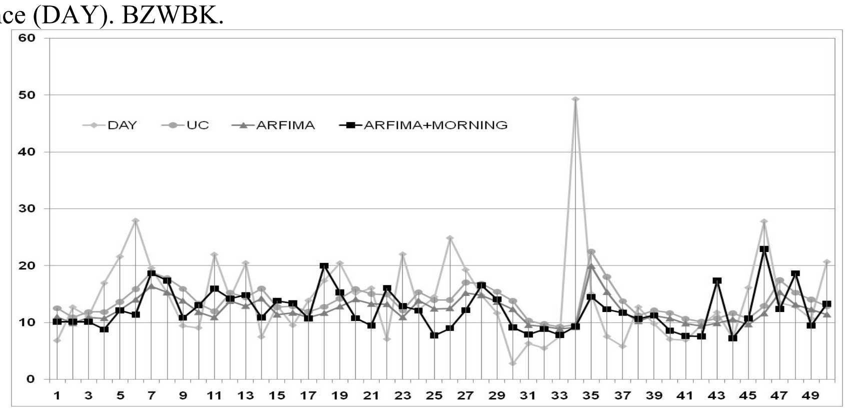
Figure 2. Forecasts obtained by means of the considered models and the daily realized variance (DAY). AGORA.



Source: Own elaboration.

The next step of the presented analysis is to compare the forecasting performance of the considered models. In each case we calculate 50 one-day-ahead forecasts. The most representative results are shown in Figures 3-4. Formal evaluation was carried out with a set of forecast quality measures. It turned out that for all the considered stocks the best results were obtained mostly by means of ARFIMA+MORNING models.

Figure 3. Forecasts obtained by means of the considered models and the daily realized variance (DAY). BZWBK.



Source: Own elaboration.

The final question we aim to address is how informative each method of forecasting is in relation to the other. To make these comparisons we run so-called encompassing regressions that can test the performance of one forecasting method against the other. These regressions take the form

$$\tilde{\sigma}^2_{[DAY]t} = \alpha + \beta_1 \hat{\tilde{\sigma}}^2_{1,\,t|t-1} + \beta_2 \hat{\tilde{\sigma}}^2_{2,\,t|t-1} + u_t \tag{8}$$

where $\hat{\tilde{\sigma}}^2_{1,\,t|t-1}$ and $\hat{\tilde{\sigma}}^2_{2,\,t|t-1}$ are the forecasts based on two different forecasting approaches. The results of this investigation are as follows. In the case of AGORA and TVN, the forecasts from the ARFIMA+MORNING model are more informative than that from ARFIMA. For the two analyzed banks, the results are different. Though the variable MORNING was significant in the ARFIMA+MORNING equation, the forecasts from this model are not more informative. The only case where the information from ARFIMA and UC models is complementary (both betas significant) is AGORA. For the other companies, the information contain of the forecasts obtained by means of these two models is similar (both betas insignificant). Taking into account UC and ARFIMA+MORNING models, we find that the results are ambiguous. For the two media companies, the information apparently is complementary. In the case of BZWBK, the forecast informativeness of ARFIMA+MORNING is higher but in the case of PEKAO both betas are insignificant.

CONCLUSIONS

The presented research addresses the question of to what extent intraday data FROM THE BEGINNING OF THE DAY CAN EXPLAIN AND PREDICT THE END-OF-THE-DAY VOLATILITY. Using the realized variance as volatility measure we show that the realized variance calculated for first few hours of the trading day (the MORNING realized variance) is a useful predictor of the daily volatility. To get possibility of some comparison, we apply 3 methods of forecasting the daily realized variance of stock returns. These are: a commonly used approach by a direct modeling as an ARFIMA process, by modeling as an ARFIMA process with the MORNING realized variance as an additional explanatory variable, and a more sophisticated way, by a direct modeling using an unobserved component model. The analysis is performed for four chosen stocks from the Warsaw Stock Exchange (AGORA, TVN, BZWBK and PEKAO). Our results show that the models taking into account the morning realized variance provide the best forecasts of the daily volatility and thus can be a useful tool in short term investments.

REFERENCES

Andersen T.G., Bollerslev T. (1998), Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts, International Economic Review 39, 885- 905

Andersen T.G., Bollerslev T., Diebold F.X., Ebens H. (2001) The Distribution of Realized Stock Return Volatility, Journal of Financial Economics, 61, 43-76

Barndorff-Nielsen O.E., Shephard N. (2002) Econometric analysis of realized volatility and its use in estimating stochastic volatility models, Journal of the Royal Statistical Society, Series B, 63, 253-280

Doman M. (2006), Modeling the Realized Volatility with ARFIMA and Unobserved Component Models: Results from the Polish Financial markets, in: Milo W., Wdowiński P. (eds), Financial Markets. Principal of Modeling, Forecasting and Decision-Making, Lodz University Press, Lodz, 123-137

Doman M., Doman R. (2004) Ekonometryczne modelowanie dynamiki polskiego rynku finansowego, Wydawnictwo Akademii Ekonomicznej w Poznaniu, Poznań

Durbin J., Koopman S.J. (2002), Time Series Analysis by State Space Methods, Oxford University Press, Oxford

Frijns B., Margaritis D. (2008), Forecasting daily volatility with intraday data, The European Journal of Finance 14, 523-540

Granger C. W. J. and R. Joyeux (1980) An Introduction to Long-Memory Time Series Models and Fractional integration, Journal of Time Series Analysis 1, 15-29

Hosking J. R. M. (1981) Fractional Differencing, Biometrica 68, 165-176

Koopman S.J., Jungbacker B., Hol E. (2005) Forecasting daily variability of the S&P 100 stock index using historical, realized and implied volatility measurement, Journal of Empirical Finance 12, 445-475

Poon S.-H. (2005) A Practical Guide to Forecasting Financial Market Volatility, John Wiley, Chichester

Tsay R.S. (2002), Analysis of Financial Time Series, John Wiley, New York

Prognozowanie zmienności zrealizowanej na koniec dnia

Streszczenie: W krótkim okresie po otwarciu giełdy ceny akcji kształtowane są przez duży pakiet nagromadzonych informacji, a zmienność zrealizowana z początku dnia ma silny wpływ na dzienną zmienność zrealizowaną. W związku z tym powstaje pytanie, czy poranna zmienność zrealizowana, wyliczana jako suma kwadratów śróddziennych stóp zwrotu, z kilku pierwszych godzin dnia giełdowego może być użyteczna w prognozowaniu dziennej zmienności zrealizowanej (zmienności na koniec dnia). W pracy stosowane są trzy różne metody prognozowania dziennej zmienności zrealizowanej akcji notowanych na Giełdzie Papierów Wartościowych w Warszawie. Uzyskane wyniki pokazują, że poranna zmienność zrealizowana może dostarczać informacji zwiększających skuteczność prognozowania dziennej zmienności zrealizowanej.

Słowa kluczowe: zmienność zrealizowana, prognozy, ARFIMA, modele składowej nieobserwowalnej